(6 C)

RECEIVED
NOV 1 3 2001
CENTER 1600/2900

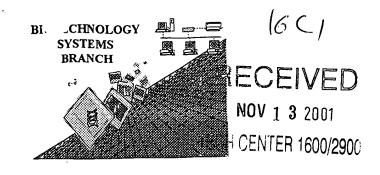# RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information
Center (STIC) detected errors when processing the following computer readable
form:

12

Application Serial Number: 09/441,242

Source: 01PE

Date Processed by STIC: 7/26/2001

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE
    APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A
    NOTICE TO COMPLY
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER
VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND
TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0
The Checker Version 3.0 application is a state-of the-art Windows based software program
employing a logical and intuitive user-interface to check whether a sequence listing is in
compliance with format and content rules. Checker Version 3.0 works for sequence listings
generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old
rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual
Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-
compliant. Checker allows public users to check sequence listings in Computer Readable form
(CRF) before submitting them to the United States Patent and Trademark Office (USPTO).
Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence
listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:
http://www.uspto.gov/web/offices/pac/checker

```
163 <210> SEQ ID NO: 4
164 <211> LENGTH: 108
165 <212> TYPE: PRT
166 <213> ORGANISM: Artificial
168 <220> FEATURE:
169 <223> OTHER INFORMATION: MTCP1 protein
171 <220> FEATURE:
172 <221> NAME/KEY: site
173 <222> LOCATION: (1)..(108)
174 <223> OTHER INFORMATION: Xaa = any amino acid
177 <400> SEQUENCE: 4
179 Met Ala Gly Glu Asp Val Gly Ala Pro Pro Asp His Leu Trp Val His
180 1               5                   10                  15
182 Gln Glu Gly Ile Tyr Arg Asp Glu Tyr Gln Arg Thr Trp Val Ala Val
183             20                  25                  30
185 Val Glu Glu Glu Thr Ser Phe Leu Arg Ala Arg Val Gln Gln Ile Gln
186             35                  40                  45
188 Val Pro Leu Gly Asp Ala Ala Arg Pro Ser His Leu Leu Thr Ser Gln
189     50                  55                  60
191 Leu Pro Leu Met Trp Gln Leu Tyr Pro Glu Glu Arg Tyr Met Asp Asn
192 65                  70                  75                  80
194 Asn Ser Arg Leu Trp Gln Ile Gln His His Leu Met Val Arg Gly Val
195                 85                  90                  95
197 Gln Glu Leu Leu Leu Lys Leu Leu Pro Asp Asp Xaa
198             100                 105
200 <210> SEQ ID NO: 5
201 <211> LENGTH: 4922
202 <212> TYPE: DNA
203 <213> ORGANISM: Artificial
205 <220> FEATURE:
206 <223> OTHER INFORMATION: genomic DNA of TCL-1
208 <220> FEATURE:
209 <221> NAME/KEY: modified_base
210 <222> LOCATION: (1)..(4922)
211 <223> OTHER INFORMATION: m = a or c
212         r = a or g
213         w = a or t
214         s = c or g
217 <220> FEATURE:
218 <221> NAME/KEY: modified_base
219 <222> LOCATION: (1)..(4922)
220 <223> OTHER INFORMATION: y = c or t
221         k = g or t
222         v = a, c or g
223         h = a, c or t
226 <220> FEATURE:
227 <221> NAME/KEY: modified_base
228 <222> LOCATION: (1)..(4922)
229 <223> OTHER INFORMATION: d = a, g or t
```

RAW SEQUENCE LISTING   DATE: 07/26/2001
PATENT APPLICATION: US/09/441,242   TIME: 15:22:26

Input Set : A:\8666-008seqlist.txt
Output Set: N:\CRF3\07262001\I441242.raw

```
    230       b = c, g, or t
    231       n = a, c, g, or t
    234  <400> SEQUENCE: 5
    235  gtcgacttgt gaktyccmag magaggccca gaagtcccgg tccggcaaag cggaggggaa    60
    237  gcggggggggg tcttccaaga agaagaaagg gcccaaggtt caacccccgg tgccttctcc   120
    239  agcagcaagc aagagggcgg cgggtcggtt gtcgctgctg gccggggccc tccgaggaaa   180
    241  ggcgcggrcc agctggggcc gggtctgcgt tcccaggagc tgccaccgtt ccagggagca   240
    243  agtcaggccg ggacgttagc gcctgcgcgg gaccctcact tgccaccaag rmccccacaa   300
    245  accccgcccc atcctgyctt acgccccgcc ccaaggtcgg ttctccccga cccgggggtc   360
W--> 247  ccgcccccaa ggnccgtcct ccccgcccce gccgsttggt ggcggccgca tgctgcccgg   420
    249  atataaaggg tcggccccac atcccaggga ccagcgagcg gccttgagag gctctggctc   480
    251  ttgcttctta ggcggcccga ggacgccatg gccgagtgcc cgacactcgg ggaggcagtc   540
    253  accgaccacc cggaccgcct gtgggcctgg gagaagttcg tgtatttgga cgagaagcag   600
    255  macgcctgcc tgcccttaac catcgaggta caaccacctt tggagcggat ggcgargcag   660
W--> 257  caggggcasc ccctgggagc ttgggatncc ctaggaaggg cgaggactca aggagcactc   720
    259  actatggggc agggaggatc ccccacagat kaagccactt ttggagccgg sctctkgagg   780
    261  gatgaatagg agttcctcca ggcagggaag aagggtggga aaaccccaaa ggaatgtcgg   840
    263  tcaaaggggt ggacccagtg cctgtggagt gtgactataa tgttgactac agcaggcatt   900
    265  ttctgggctt cggggtccta atccttaaaa atgggtatct ctaagtgact catccatatg   960
    267  gccgattatc ggaatcatct caggtgggtc ccagaaatct gtatttttaa aaagaacccw  1020
    269  cmacagttta gggtccaacc caggcataac caaaacactg gcctaagagt tgtgaagtat  1080
    271  tttcccacct accctctggg ctttatttaa gamaaccaaa tttaacaagt gatgtcgtag  1140
W--> 273  tataagcgcc ggtantkgaa ycaatattga cttttttaat gtgtgatgcc ttaagatggg  1200
W--> 275  tccttaatcc atgttaagnt tttgttaaag aaatagataa gtctttaca agtatttgga  1260
W--> 277  tttactcaat gaaaaagagt canaaaatgt tcaaactctc tccaaacata cactgaagaa  1320
W--> 279  agcataaaaa ttannaaata tattagaaca cgtatgtcca gtagcaawca maaattattg  1380
    281  agtgttgayt gtgtctctac agatgggaaa ctgaggcaca cmaaatgtac atttgtccga  1440
    283  ggtaagattg ctagtaggta atggggttgg aattctaggc tcttaaccac cacaaaatct  1500
    285  gcatttttat tggcatttca attttttaaa tatgtttta ctttaaaaat caagttaaat  1560
    287  acttactttt ttaaaatcaa aatttgaaga aataatttga agattcagtg gatttctttt  1620
    289  tttaaatctc tgagaaatct cttccctyca acgtgacacc maaaccmgcg aaccagacag  1680
    291  ttttttcataa aatcatgaaa catgcyccmc maaaaataac ccactascaa actgtgggac  1740
    293  agattttgcc tcacatcatt gaaaaggcca gcawtctttt tctctctttc tttctttgkt  1800
    295  gttttttttt tttcctgtag awacagggtc tcgctctgtg acccaggctg gtctyaaact  1860
    297  cctggcctca agcgatcctc ctgcctctgc cttccaaagc actggaatta caagtgtgag  1920
    299  ccgctgcaac ccgccagaaa aaagtgtgcc tttcatggcc ctgtctgggt ggctagacac  1980
    301  gtgtgtgtgc tggtggtcct ggcccagcca gagttccctg agaggagcat gcatggccta  2040
    303  aggaagtgag cttcagggaa cagtgatgac catcatttca cactcggacc ccctgccmaa  2100
W--> 305  gatgggtgga tgsctgscag ggagggattc cggtkttcct gcgcctggag aanccctgcc  2160
W--> 307  aagcggaacc tgaaagtatn ccctgtcctt ttcttctcct nagataaagg ataggttaca  2220
W--> 309  gttnngggtg ctcttgcgtc gggaagacgt cgtcctgggg aggcctatga cccccaccna  2280
    311  gataggccca agcctgctgc ctatcatgtg gcagctctac cctgatggac gataccgatc  2340
    313  ctcagactcc agtttctggc gcttagtgta ccacatcaag gtgagtgtct ttctcccaga  2400
    315  ggtccatcgg ktgatcttgg gtttcccctc cycmatgtct gsccttagtg gtttaycttc  2460
W--> 317  ccyccatccc agtssgcaaa agcattwaaa aratggggga nrtrwacmas tgcagatttc  2520
W--> 319  tanaggactt taccagagag aaganagatc ctntgaggtc tctaanagaa ccctacctcc  2580
W--> 321  acttcctccc anccaccanc taccgcagg aagacatctc tggtggggmm kcacaggctg  2640
    323  aaggctggtg ggaggagggr caktctccaa gasccctga aatcctcaca cctgggttcc  2700
    325  tacctgctgt ttccagctag gggaagscsc aggagtgagg aatggaggga gtggagggct  2760
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/441,242

DATE: 07/26/2001
TIME: 15:22:26

Input Set : A:\8666-008seqlist.txt
Output Set: N:\CRF3\07262001\I441242.raw

```
      327 ctggccgatc aatgccttct ctctctctct gcctctcaga ttgacggcgt ggaggacatg   2820
      329 cttctcgagc tgctgccaga tgactgatgt atggtgagct ccactggagc ctgacccctc   2880
      331 ttagtccaca gtggctgtat cagaaagaaa gaccacccct tctccatgaa ggcagtgcta   2940
      333 acccctcccc gactgctgcc atctgagggt ccctagggat gggagaggct tcctggaggc   3000
      335 actcatgtct cccttaccac ttcgggagcc aagggctttg gtaggcagcc ccctttatcg   3060
      337 cagctgctca tatctataaa gtacttcaca agtttcagct ggcactttca ttttaccatt   3120
      339 gctttttttt tctttgggag atgagtctgg ctctgtggcc caggctagag tgtagtgggt   3180
      341 gcaatctcag ctcactgaaa gctctgcctc ccgggttcac accattctcc tgcctcagcc   3240
      343 ctcggagtag ctgggactac aggcgcccgc caccacacct ggctaatttt ttttttttw    3300
      345 ttwtwtttt tagtagagmc ggggtttcac cgtgttagcc aggatggtct cgatctcctg    3360
      347 acctcatgat ctgcccgcct cggcctccca aagtgctggg attacaggca tgagccacca   3420
      349 cgtccggcct taccattgct ttattaaata agcactggtg cttgattata tcagctgagc   3480
      351 cagatattag atacgctatt gagttttgrg gaaataagag taccaaaact cagaaatgag   3540
      353 ttgaagtata gtgacatctt cagattacag acccaggtgt cagaatttgc cttggctcag   3600
      355 aaggcctctg ggggccatcc ctgaccacta ggcttccac ttagacctgc tccagcagca    3660
      357 ccacccctcg scactgcctg gtcctttcct tcacccttga ttctgtcttc ttttgtcctt   3720
      359 ctccaggtct tggyagcacc tgtctccttt caccccaggg cctgagcctg gccagcctac   3780
      361 aatggggatg ttgtgtttct gttcaccttc gtttactatg bctgtgtctt ctccaccacg   3840
      363 ctggggtctg ggaggaatgg acagacagag gatgagctct acccrgggcc tgsaggacct   3900
W--> 365 gtcctgtagm ccactctgct cgccttagsa cctacsactc cwrccgasga ggatnccant   3960
W--> 367 tggaagagct tcttnnaggt gncnaanaan anctgtgcgt nggcttttct cagctggatg   4020
W--> 369 atggtcntna gcctctttct gtcccttctg tccctcacag cactagtatt tnatgttgca   4080
      371 cacccactca gctccgtgaa tttgtgagaa cacaaccgat tcacctgagc aggacctctg   4140
W--> 373 aaaccctgga ccagtggtct cacatggtgc tacgcctgca tgtaaacacg cctncaaacg   4200
      375 ctgcctgcck gtraacacgm sksyrmacag stgmswrccc gtaaacacgc ctgcaaacgc   4260
      377 tgcctgccca cacaggttca cgtgcagctc aaggaaagrm ctgaaarrag cccttatctg   4320
W--> 379 tgctcaggac tcagaagcct ctgggtcagt ggtccacatc ccgggacgca gnaggaggcc   4380
      381 aggccggcga gccctgtgga tgagccctca gaacccttgg gttgcccacg tggaaaaggg   4440
      383 atagaggttg ggtttccccc cttttataga tggtcacgca cctgggtgtt acaaagttgt   4500
W--> 385 atgtggcatg aatacttgnt gtnatgattg attaaatgca agatagttta tctaacttcg   4560
W--> 387 tgcggaatca gcttctatcc ttgncttaga ttctggtgga gagaagtgan aataggcagn   4620
W--> 389 ccccanataa anaatattca ngggatttat tttattnttc cttttggggng atnngggact   4680
W--> 391 acattntncn nccccgtnta atccaatgnt taaanccca gtgttcttgg aggncntacg     4740
W--> 393 tcgaanacca ttggngtang caacctcaaa atttttnngt tgnnaattnc cngttttcca   4800
W--> 395 gagncccccc cntnctccat cttnntcctn gcccnccctn ncctcccnca ntcccnangt   4860
W--> 397 tncccntcgnc cccagtcagt tctttctccn ncttnanccg ntnatntcac cagnttcttt   4920
      399 ct                                                                   4922
      402 <210> SEQ ID NO: 6
      403 <211> LENGTH: 20
      404 <212> TYPE: DNA
      405 <213> ORGANISM: Artificial
      407 <220> FEATURE:
      408 <223> OTHER INFORMATION: p9A primer
      410 <400> SEQUENCE: 6
      411 tgctgccaga tgactgatgt                                                  20
      414 <210> SEQ ID NO: 7
      415 <211> LENGTH: 20
      416 <212> TYPE: DNA
      417 <213> ORGANISM: Artificial
```

Use of n and/or Xaa has been detected in the Sequence Listing. Review the Sequence Listing to insure a corresponding explanation is presented in the <220> to <223> fields of each sequence using n or Xaa.

VERIFICATION SUMMARY   DATE: 07/26/2001
PATENT APPLICATION: US/09/441,242   TIME: 15:22:27

Input Set : A:\8666-008seqlist.txt
Output Set: N:\CRF3\07262001\I441242.raw

L:19 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:1
L:97 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:2
L:99 M:258 W: Mandatory Feature missing, <220> FEATURE:
L:99 M:258 W: Mandatory Feature missing, <223> OTHER INFORMATION:
L:136 M:220 C:. Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:3
L:166 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:4
L:197 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:203 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:5
L:247 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:257 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:273 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:275 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:277 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:279 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:305 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:307 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:309 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:317 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:319 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:321 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:365 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:367 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:369 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:373 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:379 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:385 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:387 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:389 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:391 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:393 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:395 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:397 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:405 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:6
L:417 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:7
L:429 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:8
L:441 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:9
L:453 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:10
L:465 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:11
L:477 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:12
L:511 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12

ERROR DETECTED   **SUGGESTED CORRECTION**   SERIAL NUMBER: 09/441,242

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1____Wrapped Nucleics Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
2____Invalid Line Length	The rules require that a line not exceed 72 characters in length. This includes white spaces.
3____Misaligned Amino Numbering	The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
4____Non-ASCII	The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
5____Variable Length	Sequence(s)____contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
6____PatentIn 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s)_____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
7____Skipped Sequences (OLD RULES)	Sequence(s)_____ missing. If intentional, please insert the following lines for each skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
8____Skipped Sequences (NEW RULES)	Sequence(s)_____ missing. If intentional, please insert the following lines for each skipped sequence. <210> sequence id number <400> sequence id number 000
9____Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
10____Invalid <213> Response	Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
11____Use of <220>	Sequence(s)__2____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
12____PatentIn 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
13____Misuse of n	n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

OIPE

```
         RAW SEQUENCE LISTING                DATE: 07/26/2001
         PATENT APPLICATION:  US/09/441,242  TIME: 15:22:26

         Input Set : A:\8666-008seqlist.txt
         Output Set: N:\CRF3\07262001\I441242.raw
```

```
 3 <110> APPLICANT: Russo, Giandomenico et al.
 5 <120> TITLE OF INVENTION: TCl-1 Gene and Protein and Related Methods and Compositions
 7 <130> FILE REFERENCE: 8666-008
 9 <140> CURRENT APPLICATION NUMBER: 09/441,242
10 <141> CURRENT FILING DATE: 1999-11-16
12 <160> NUMBER OF SEQ ID NOS: 12
14 <170> SOFTWARE: PatentIn version 3.0
16 <210> SEQ ID NO: 1
17 <211> LENGTH: 1324
18 <212> TYPE: DNA
19 <213> ORGANISM: Artificial
21 <220> FEATURE:
22 <223> OTHER INFORMATION: cDNA sequence of TCL-1
24 <220> FEATURE:
25 <221> NAME/KEY: CDS
26 <222> LOCATION: (49)..(387)
28 <400> SEQUENCE: 1
29 cttgagaggc tctggctctt gcttcttagg cggcccgagg acgccatg gcc gag tgc       57
30                                                      Ala Glu Cys
31                                                      1
33 ccg aca ctc ggg gag gca gtc acc gac cac ccg gac cgc ctg tgg gcc      105
34 Pro Thr Leu Gly Glu Ala Val Thr Asp His Pro Asp Arg Leu Trp Ala
35     5               10              15
37 tgg gag aag ttc gtg tat ttg gac gag aag cag cac gcc tgg ctg ccc      153
38 Trp Glu Lys Phe Val Tyr Leu Asp Glu Lys Gln His Ala Trp Leu Pro
39 20              25              30              35
41 tta acc atc gag ata aag gat agg tta cag tta cgg gtg ctc ttg cgt      201
42 Leu Thr Ile Glu Ile Lys Asp Arg Leu Gln Leu Arg Val Leu Leu Arg
43             40              45              50
45 cgg gaa gac gtc gtc ctg ggg agg cct atg acc ccc acc cag ata ggc      249
46 Arg Glu Asp Val Val Leu Gly Arg Pro Met Thr Pro Thr Gln Ile Gly
47             55              60              65
49 cca agc ctg ctg cct atc atg tgg cag ctc tac cct gat gga cga tac      297
50 Pro Ser Leu Leu Pro Ile Met Trp Gln Leu Tyr Pro Asp Gly Arg Tyr
51         70              75              80
53 cga tcc tca gac tcc agt ttc tgg cgc tta gtg tac cac atc aag att      345
54 Arg Ser Ser Asp Ser Ser Phe Trp Arg Leu Val Tyr His Ile Lys Ile
55     85              90              95
57 gac ggc gtg gag gac atg ctt ctc gag ctg ctg cca gat gac             387
58 Asp Gly Val Glu Asp Met Leu Leu Glu Leu Leu Pro Asp Asp
59 100             105             110
61 tgatgtatgg tcttggcagc acctgtctcc tttcacccca gggcctgagc ctggccagcc      447
63 tacaatgggg atgttgtgtt tctgttcacc ttcgtttact atgcctgtgt cttctccacc      507
65 acgctggggt ctgggaggaa tggacagaca gaggatgagc tctacccagg gcctgcagga      567
67 cctgcctgta gcccactctg ctcgccttag cactaccact cctgccaagg aggattccat      627
69 ttggcagagc ttcttccagg tgcccagcta tacctgtgcc tcggcttttc tcagctggat      687
71 gatggtcttc agcctctttc tgtcccttct gtccctcaca gcactagtat ttcatgttgc      747
```

```
 73 acacccactc agctccgtga acttgtgaga acacagccga ttcacctgag caggacctct    807
 75 gaaaccctgg accagtggtc tcacatggtg ctacgcctgc atgtaaacac gcctgcaaac    867
 77 gctgcctgcc ggtaaacacg cctgcaaacg ctgcctgccc gtaaacacgc ctgcaaacgc    927
 79 tgcctgccca cacaggttca cgtgcagctc aaggaaaggc ctgaaaggag cccttatctg    987
 81 tgctcaggac tcagaagcct ctgggtcagt ggtccacatc ccgggacgca gcaggaggcc   1047
 83 aggccggcga gccctgtgga tgagccctca gaacccttgg cttgcccacg tggaaaaggg   1107
 85 atagaggttg ggtttccccc ctttatagat ggtcacgcac ctgggtgtta caaagttgta   1167
 87 tgtggcatga atactttttg taatgattga ttaaatgcaa gatagtttat ctaacttcgt   1227
 89 gcgcaatcag cttctatcct tgacttagat tctggtggag agaagtgaga ataggcagcc   1287
 91 cccaaataaa aaatattcat ggaaaaaaaa aaaaaaa                             1324
 94 <210> SEQ ID NO: 2
 95 <211> LENGTH: 113
 96 <212> TYPE: PRT
```

C--> 97 <213> ORGANISM: (Artificial) → see item 11 on Exa Summary Sheet

W--> 99 (<220>) FEATURE:

W--> 99 (<223>) OTHER INFORMATION:

```
 99 <400> SEQUENCE: 2
101 Ala Glu Cys Pro Thr Leu Gly Glu Ala Val Thr Asp His Pro Asp Arg
102 1               5                   10                  15
105 Leu Trp Ala Trp Glu Lys Phe Val Tyr Leu Asp Glu Lys Gln His Ala
106             20                  25                  30
109 Trp Leu Pro Leu Thr Ile Glu Ile Lys Asp Arg Leu Gln Leu Arg Val
110             35                  40                  45
113 Leu Leu Arg Arg Glu Asp Val Val Leu Gly Arg Pro Met Thr Pro Thr
114       50                  55                  60
117 Gln Ile Gly Pro Ser Leu Leu Pro Ile Met Trp Gln Leu Tyr Pro Asp
118 65                  70                  75                  80
121 Gly Arg Tyr Arg Ser Ser Asp Ser Ser Phe Trp Arg Leu Val Tyr His
122             85                  90                  95
125 Ile Lys Ile Asp Gly Val Glu Asp Met Leu Leu Glu Leu Leu Pro Asp
126             100                 105                 110
129 Asp
133 <210> SEQ ID NO: 3
134 <211> LENGTH: 560
135 <212> TYPE: DNA
136 <213> ORGANISM: Artificial
138 <220> FEATURE:
139 <223> OTHER INFORMATION: genomic sequence of TCL-1
141 <400> SEQUENCE: 3
142 gtcgactgtg agttcccagc agaggcccag agtcccggtc cggcagccga gggaagcggg     60
144 ggggtcttcc agaagaagaa agggccaagg tcaccccggt gcctctccag cagcagcaga    120
146 gggcggcggt cggtgtcgct gctggccggg gcctcgagga aggcgcgggc cagctggggc    180
148 cgggtctgcg ttcccaggag ctgccaccgt tccagggagc aagtcaggcc gggacgttag    240
150 cgcctgcgcg ggaccctcac ttgccaccaa ggaccccaca aaccccgccc catccttagc    300
152 gcctgcgcgg gaccctcact tgccaccaag accccacaa accccgcccc atcctgcctt    360
154 acgccccgcc ccaaggtcgt tctcccgacc cggggtcccg ccccaagacc gtcctcccgc    420
156 cccgccgctt ggtggcggcc gcatgctgcc cggatataaa gggtcggccc cacatcccag    480
158 ggaccagcga gcggccttga gaggctctgg ctcttgcttc ttaggcggcc cgaggacgcc    540
160 atggccgagt gcccgacact                                                560
```